DOCUMENT RESUME

ED 209 357 TM 810 933

AUTHOR	Brougham, James
TITLE	The Measurement of Language Diversity.
INSTITUTION	Laval Univ., Quebec (Quebec). International Center for Research on Bilingualism.
REPORT NO	ICRB-B-98
PUB DATE	81
NOTE	36p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Bilingual Education; *Bilingualism; Language Proficiency; Languages for Special Purposes; Language Standardization; Language Usage; *Measurement Techniques; Second Language Learning; Sociolinguistics
IDENTIFIERS	*Language Diversity

ABSTRACT

Accepting that language diversity is functionally related to other variables characterizing human societies, much discussion stems from the advantages or disadvantageous nature of language diversity in terms of national development and national unity. To discover ways of measuring language diversity would help, in part, to solve the language diversity issue; however, the lack of consistency and agreement in the definition of the two viewpoints hampers the language planners. Bearing in mind that any language diversity measure takes into consideration all languages present and considers the numbers of users of the languages, the coupling of these two independent variables renders the elimination of all ambiguity impossible in a diversity measure. For example, a society bearing a large number of languages with widely differing numbers of users will have the same diversity measurement as one characterized by a smaller number of languages but greater evenness of user distribution. The diversity measurement must be related to both the number of languages and the degree of evenness of user distribution. Using indices and equation models of the research of language planners, this document sets out to define properties which a language diversity index should exhibit, using both sample and census data. (Author/CE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

centre international de recherche sur le bilinguisme

ED209357

publication
B-98

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

A. Prujiner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

THE MEASUREMENT OF LANGUAGE DIVERSITY

TM 810 932

James Brougham

1981

CIRB
ICRB

James Brougham

The Measurement of Language Diversity

Publication B-98

1981
Centre international de recherche sur le bilinguisme
International Center for Research on Bilingualism
Québec

Le Centre international de recherche sur le bilinguisme est un organisme de recherche universitaire qui reçoit une subvention de soutien du ministère de l'Éducation du Québec et une contribution du Secrétariat d'État du Canada pour son programme de publication.

The International Center for Research on Bilingualism is a university research institution which receives a supporting grant from the Department of Education of Quebec and a contribution from the Secretary of State of Canada for its publication programme.

© 1981 CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME
Tous droits réservés. Imprimé au Canada
Dépôt légal (Québec) : 2e trimestre 1981

THE MEASUREMENT OF LANGUAGE DIVERSITY

INTRODUCTION

If is frequently asserted that language diversity is functionally related to other variables characterising human societies. (See, for instance, Greenberg 1956, Gumperz 1962, Fishman 1968a, 1968b, 1977, Pool 1969, Lieberson & Hansen 1974). A perennial bone of contention, for example, has been the advantageous or disadvantageous nature of language diversity in terms of national development and national unity (Fishman 1963, Deutsch 1966, Siméon 1972). As Pool (1969) has pointed out, however, the current ability of language planners to estimate the relevance of language diversity to development is almost nil, and one of the main reasons for this is the lack of consistency and agreement in the definition of the two concepts involved. The purpose of the present article is not to add to the development/diversity debate, but rather to suggest reasonable and internally consistent ways of measuring language diversity. It is hoped that this contribution will enable ensuing discussions of the relationship between diversity and other variables to be pursued upon the basis of less ambigous, comparable evaluations of this elusive phenomenon.

Desirable Properties of a Language Diversity Index

Although Pool (1969) emphasises the wooliness of contemporary definitions of the concept of language diversity, he continues to employ an apparently arbitrary definition in his attempt to clarify the relationship between this variable and national development. The measure used in his article was the size of the « largest native language community (% of population) »[1] and variations on this theme have frequently appeared in the literature (Banks & Textor, 1963, Fishman, 1968, Fishman, Cooper and Rosenbaum, 1977, Criper and Ladefoged, 1971). Although the most straightforward and unambiguous way of measuring language diversity is simply the number of languages[2] coexisting within a given observational unit, almost all writers have in fact sought to incorporate numbers or proportions of language users into their diversity measures. The general consensus of opinion would appear to be that the more even the distribution of users among the language categories the greater the language diversity of the unit in question. Some writers, however, characterise the units under observation as being diverse or homogeneous without regard to the numbers of different languages in use within them. Banks and Textor, (1963) for instance, define polities in which one language is natively spoken by 85% of the population and in which no significant linguistic minority is present as being less diverse than polities in which the remaining 15% of speakers may be assigned predominantly to a single lan-

[1] Pool, p. 148.

[2] Throughout this article it is assumed that a discrete classification of languages is feasible.

guage[3]. This can give rise to considerable ambiguity in that the members of the former class are actually likely to exhibit a larger number of languages and yet be classified as less diverse.

It is the opinion of the present writer that any language diversity measure must take into account first and foremost *all* the languages present but should also seek to incorporate the legitimate concern expressed by most writers for the numbers of users of these languages. Although it is not anticipated that this goal will be disputed, it should be borne in mind that the coupling of two independent variables, namely number of languages and the distribution of speakers among them, renders the elimination of *all* ambiguity impossible in a diversity measure. It is quite possible, for instance, that a society featuring a large number of languages, with widely differing numbers of users will have the same diversity measurement as a society characterised by a smaller number of languages but greater evenness of user distribution. Any diversity measurement should therefore be evaluated in the light of both the number of languages and the degree of evenness of user distribution (see p. 6).

The desire to incorporate numbers and distributions of speakers in a diversity index calls for another caveat at this point. As Fishman (1968b) has pointed out, the demographic status of a language does not necessarily coincide with its use and functions within a given society. A fundamental aspect of any demographically based measure of diversity is, therefore, that it constitutes a surrogate measure of the diversity of a limited subset of the total set of linguistic *activities*. The type of linguistic activity whose diversity it is required to measure should be clearly specified at the outset. It is for this reason that the more general locution 'language users' is prefered to that of 'language speakers' throughout this theoretical paper. A language may be widely read and therefore used in the context of work, for instance, without being spoken. Indices based on numbers or proportions of native speakers, on the other hand, tell us something about the diversity of languages spoken in the homes of a given census or survey unit at sometime in the past, (providing migration is allowed for). They cannot be taken to describe diversity of actual linguistic *activity* outside the home at a later point in time, as Lieberson (1968), for example, seems to assume in his development of Greenberg's diversity index. As mother tongue data are by far the most commonly used for diversity measurement purposes, it must be further emphasised that, since home languages (and therefore mother tongues) vary over time, evaluations of diversity rooted in demographic data should not be based on widely differing age groups unless one is simply interested in measuring the mother tongue diversity of a set of human beings *per se*. An example may help to make this clear. A mother tongue based measure of the diversity of the population in the North West Highlands of Scotland, which did not control for age, would produce a reasonably high diversity value. It would be erroneous, however, to conclude that this represents the diversity of linguistic *activity* in the region's homes at any one point in time. For the older members of the population (≥60) the home language was almost exclusively Gaelic whereas for the younger generation (≤20), English has been more or less unchallenged as the language of the home. The degree of mother tongue diversity among the population as a whole at the present time is undoubtedly higher than that of linguistic *activity* among the region's homes at present or indeed at most periods in the past. Mother tongue diversity then can only be used as a surrogate

[3] Lieberson (1975) has shown that the size of the largest mother tongue group is in fact an excellent nonlinear predictor of one of the more satisfactory measures of diversity, Greenberg's A-index, defined below.

measure for *home* linguistic activity within a societal unit, and should ideally be calculated using data in which age has been controlled for (age cohorts).

As stated on page (2) the inclusion of both numbers of languages and the distribution of users of these languages in a diversity measure will inevitably entail a certain degree of subjectivity, as the calculated values will depend upon the definition of the relationship between these two independent variables. Hurlbert (1971) gives some examples of these in his critique of the diversity concept. Nevertheless, certain properties of such a measure are clearly desirable and here much can be learned from the work of ecologists since the Second World War. Quantitative ecologists have devoted considerable effort to the development and/or application of diversity indices to the species make-up of natural communities, often with a view to relating species diversity to other community properties such as productivity and stability (Pielou 1967). Such concerns are quite analogous to the types of problem to which Fishman (1968a), Pool (1969), Lieberson (1974, 1975) and other scholars have adressed themselves. While no-one would suggest that languages obey the laws of biology, the logical problem of measuring diversity is identical in both cases. This is exemplified by the fact that the only well known measure of language diversity which incorporates both numbers of languages and corresponding numbers of users, namely, Greenberg's A-index (Greenberg, 1956), is formally almost identical to one of the more widely used ecological indices of diversity, Simpson's index (Simpson, 1949), although both appear to have been developed independently (see below, p. 7).

Pielou (1975) lists three desirable properties for a diversity index, D, which is to be a function of both the numbers of categories (languages) and the relative frequencies of items within those categories (proportions of users), D (p_1, p_2, ... p_ℓ). Weinreich (1957) has also evoked the first two of these properties while Greenberg (1956) has indicated the desirability of property (3).

Expressed in sociolinguistic terms, these may be read as :

Property 1

For any given ℓ (number of languages), D should take on its greatest value when $p_i = 1/\ell$ for all i, i.e. when users are apportioned evenly among all the languages present in the societal unit under consideration (*Note the subjectivity of the relationship between these two variables mentioned earlier; one could conceivably define diversity in exactly the opposite way*).

Property 2

Given two societal units in which users are apportioned evenly among languages, one with ℓ languages and one with $\ell + 1$ languages, then D should take on a greater value in the latter case.

Property 3

Given two societal units characterised by identical distributions of numbers of languages and users, then D should take on a greater value for a unit wherein the observed languages belong to different language groups than for a unit wherein the observed lan-

guages belong to a single language group, i.e. the diversity index should take into account the hierarchical nature of language classification and hence the concept of interlingual distance.

More formally, suppose that language users are subjected to two different classifications, namely language group classification, G, with g classes and a language classification, L, with ℓ classes. Let p_i ($i = 1, \ldots, g$,) be the proportion of speakers in the ith class of the G - classification and let p_{ij} ($i = 1, \ldots, g$; $j = 1, \ldots, \ell$) be the proportion of these speakers in the jth class of the L classification. Let $\pi_{ij} = p_i p_{ij}$ be the proportion of the whole community belonging to the ith G class and the jth L class.

Now, let D (GL) be the diversity of the doubly classified population; D (G) its diversity under the language group classification and D_i (L) the diversity under language classification of those speakers belonging to the ith G class. If, in addition, we let D_G (L) be the average of the D_i (L) over all G classes, it is then required that

$$D(GL) = D(G) + D_G(L)$$

In addition to allowing for a possible hierarchical classification of languages, property 3 would provide the possibility of measuring a population's diversity not only in terms of linguistic criteria but also in terms of kindred criteria or, indeed, totally unrelated criteria. One could conceive, for instance, of an ethnolinguistic diversity index based on a classification by ethnic affiliation and a classification by mother tongue.

Pielou (1969) has shown that the only function of the $(p_1, p_2, \ldots, p_\ell)$ proportions having these three properties is

$$(1) \qquad D(p_1, p_2, \ldots, p_\ell) = - C \sum_{i=1}^{\ell} p_i \log p_i$$

where C is a positive constant. If C is set equal to 1 we are left with the index

$$(2) \qquad D = - \sum_{i=1}^{\ell} p_i \log p_i$$

The diversity of a doubly classified population (property 3) would then be given by (Pielou, 1969)

$$(3) \qquad D_{(GL)} = - \sum_{i=1}^{g} p_i \log p_i + \sum_{i=1}^{g} p_i D_{i(L)}$$

$$\text{where } D_{i(L)} = - \sum_{j=1}^{\ell} p_{ij} \log p_{ij}$$

Diversity Measurements Using Sample Data

Equation (2) is usually referred to as the Shannon index (Shannon and Wiener, 1949) and is generally described as H', the measure of the information content of a code. In the language planning context it may be interpreted as the diversity per individual in a multilingual context. It should be noted that the Shannon index as defined in information theory is, strictly speaking, valid only for an infinite population. It measures the average information contained in a code in the long run rather that the information contained in a particular message. It should therefore be used only for societal units that are infinitely large in the sense that removing samples for them causes no perceptible change in them. [When complete census data are available diversity may be characterised quantitatively by using Brillouin's index (see below).]

The only use of this index bordering on an application to the measurement of language diversity, albeit in a non-sampling context, appears to be Sadler's (1962) little known paper. Sadler uses a different formulation of the Shannon index

$$\text{Log I} = \log N - \frac{1}{N} \sum_{i=1}^{K} n_i \log n_i \qquad (4)$$

where N is the sum of all items, (in his published examples individuals classified according to their nationality), and n_i is the number of items (individuals) in the ith category (nationality). The antilog of the resulting value, namely I, is then interpreted as a measure of the internationality of the organisations or sets of conference delegates with which he was concerned. « I tells us how many nations with equal representation would be equivalent to the observed distribution »[4]. While this interpretation of the resulting figure is a valid one in the context with which Sadler was dealing, it is perhaps unfortunate that it can result in higher measures of internationality (diversity) for organisations with narrower ranges of nationalities. This is because Sadler privileges the attainment of evenness of distribution of individuals at the expense of category diversity. This contrasts with the ecologists' approach, where a given diversity measurement is often compared with the maximum value which it could attain, given the same number of categories and a redistribution of the items involved. Although the definition of the relationship between categories and numbers of items inevitably involves a certain degree of subjectivity, as mentioned earlier, it is surely more reasonable to regard the actual number of categories as being the more important of the components, and to redistribute individuals among the categories, when evaluating diversity for distributions of maximum evenness with which to compare the diversity of observed distributions. Lieberson (1969) has stated that ratios between diversity measurements and their corresponding, maximum possible values (given the same number of categories) should not be used for comparisons between units featuring different number of categories. He suggests that such 'standardisation procedures' give misleading results in that the resulting 'standardised measures' may be strikingly different from those obtained using the basic diversity measures. It must be borne in mind, however, that evenness and diversity are different aspects of the partition of numbers of users among languages. Ratios of this type are not alternative diversity indices but rather measures of the evenness of the distribution of individuals among the various language categories. Inter-

[4] Sadler, p. 480.

unit comparisons should ideally resort to both indices, as mentioned on page (2) of the present article.

Maximum diversity for Shannon's index under conditions of complete evenness with ℓ languages is

$$(5) \qquad D_{max} = -\Sigma \frac{1}{\ell} \log \frac{1}{\ell} = \log \ell$$

The ratio of the observed diversity value to the maximum possible value (assuming the same number of individuals and languages) may then be taken as a measure of the degree of evenness of the apportionment of users to the various languages

$$(6) \qquad \text{Evenness} = \frac{D_{obs}}{\log \ell}$$

According to Poole (1974) equation (2) is a biased estimate of population diversity. Where the number of languages present in the unit under consideration is known (as it usually will be), the expected value of the observed diversity value, D_{obs}, is given by the series (Hutcheson 1970).

$$(7) \quad E\,(D_{obs}) = \left(-\sum_{i=1}^{\ell} p_i \ln p_i\right) - \left(\frac{\ell - 1}{2N}\right) + \left(\frac{1 - \Sigma p_i^{-1}}{12N^2}\right) + \left(\frac{\Sigma(p_i^{-1} - p_i^{-2})}{12N^3}\right) + \ldots$$

Natural logarithms must be used for the calculation of $E\,(D_{obs})$, the third and subsequent terms of which are usually very small. Hutcheson (1970) has also shown the variance of (D_{obs}) to be

$$(8) \qquad \text{Var}\ (D_{obs}) = \frac{-\sum_{i=1}^{\ell} p_i \ln^2 p_i - (\sum_{i=1}^{\ell} p_i \ln p_i)^2}{N} + \frac{\ell - 1}{2N^2} + \ldots$$

The $E\,(D_{obs})$ of two different samples can be compared by means of a 't' - test, to see, for example, if language diversity is changing over time. 't' would be given by (Hutcheson 1970)

$$(9) \qquad \frac{D_{obs_1} - D_{obs_2}}{|(\text{Var}\left(D_{obs_1}\right) - \text{Var}\left(D_{obs_2}\right)|^{1/2}}$$

and the degrees of freedom for the test by (Hutcheson 1970)

$$(10) \qquad \frac{[\mathrm{Var}\left(D_{obs_1}\right) - \mathrm{Var}\left(D_{obs_2}\right)]^2}{\left(\mathrm{VarD}_{obs_2}\right)^2/N_1 + \left(\mathrm{VarD}_{obs_2}\right)^2/N_2}$$

The usefulness of the preceding formulas to language planners in countries such as Canada, which are seeking to implement multicultural or multilingual policies is readily apparent. They permit regular monitoring of the national diversity situation and its evolution without resorting to cumbersome, full scale census taking. For a fuller discussion of the distributional properties of Shannon's and some other indices of diversity the reader is referred to the 1969 paper by Bowman, Hutcheson, Odum and Shenton.

In the context of the measurement of language diversity it is perhaps worthwhile mentioning that the Shannon diversity index is a special case of a more general class of functions used in the mathematical theory of information. Renyi (1961) has shown that, given a code of symbols, the function

$$(11) \qquad H_a = \frac{\log \Sigma\ p_i^{\,a}}{1 - a}$$

is the entropy of order a of the code. Setting a equal to 1 it can be shown (Pielou 1975) that

$$(12) \qquad \lim_{a \to 1} H_a = -\Sigma p_i \log p_i = D_{Shannon} = H_1$$

or the entropy of order 1 of the set of the p_i (proportions in category i). With a = 2 we obtain

$$(13) \qquad H_2 = -\log \Sigma p_i^{\,2}$$

The function Σp_i^2 is the only diversity measure which has so far received any widespread application in the field of sociolinguistics. Readers will recognize it as Greenberg's (1956) A-index which describes the probability that any two randomly picked individuals from a given societal unit will share the same mother tongue (or indeed any other linguistic or extra-linguistic feature).

It is interesting to note that Greenberg's index is almost identical to that proposed by Simpson (1949) as a measure of ecological diversity, namely

(14) $$D_{Simpson} = \frac{\Sigma n_i (n_i - 1)}{N (N - 1)}$$

where n_i = number of individuals in category i and N = total of individuals. The difference is that Simpson suggested choosing individuals without replacement. Lieberson (1969) has pointed out that essentially similar indices have in fact been developed by workers in a variety of other fields. As in the case of Greenberg's index, ecologists tend to follow Pielou's (1969) recommendation of subtracting the resulting values from unity in order to obtain an index which increases with diversity rather than uniformity. Thus

(15) $$D_{Simpson} = 1 - \Sigma \frac{n (n - 1)}{N (N - 1)}$$

and

(16) $$D_{Greenberg} = 1 - \Sigma \left(\frac{n_i}{N}\right)^2$$

This measure of diversity is widely used by ecologists and has received some application to census data on mother tongue diversity by Greenberg and later by Lieberson (1964, 1974, 1975a, 1975b).

Although it can be shown (Pielou 1969) that the Simpson index, and by extension the Greenberg index, are formally identical for both fully censused and sampled populations, it will be recalled that Shannon's formula (2) is inappropriate for fully censused populations (see page 5). When census data are available to the investigator an appropriate measure of diversity is Brillouin's index (Brillouin 1962). The measurement of the language diversity of fully censused societal units is discussed below.

It is perhaps not superfluous at this juncture to point out that Hill (1973) has demonstrated that different indices measure different aspects of the partition of items among categories. They differ in the importance which they assign to the rarer or more commonly used languages respectively. This corresponds to the inevitable subjectivity in diversity indices mentioned earlier. He suggests that, rather than taking the logarithms of entropies, which are 'harder to visualise', diversity numbers, defined as the reciprocal of the $(a-1)^{th}$ root of a weighted mean of the $(a-1)^{th}$ powers of the proportional abundances of the n categories should be used. More formally

(17) $$N_a = \left(p_1^a + p_2^a + \ldots + p_n^a\right)^{1/(1-a)}$$

where N_a is the diversity number of order a. N_0 would then measure diversity in terms simply of the number of languages present; N_1 would be $\exp(D_{Shannon})$ and N_2 would be the reciprocal of Greenberg's index, i.e. $1/(p_1^2 + p_2^2 + \ldots + p_n^2)$. As the order of the diversity number increases the importance assigned to the more widely used languages would augment.

Hill further claims that N_a is a strictly decreasing function of a and that N_1, although a transformation of Shannon's index, is in no way exceptional. However, as Pielou (1969, 1977) shows, the only function of the proportions of items, $(p_i, \ldots\ldots p_\varrho)$, having property (3) (see page 3) is Shannon's formula and it is therefore to be preferred to Greenberg's index as a more flexible tool for the measurement of the language diversity of sample data sets when this property is required.

Diversity Measurement Using Census Data

Brillouin's information-theoretic index is defined as (Brillouin 1962)

$$B = \frac{G\ !}{N_o\ !\ .\ N_1\ !\ \ldots .\ N_j\ !} \tag{18}$$

where G is the total number of symbols in a code and $N_o, \ldots, N_j$ are the numbers of symbols of each different kind. Insofar as the measurement of language diversity is concerned equation (18) may be reformulated as

$$D_B = \frac{1}{N} \log \frac{N\ !}{N_1\ !\ .\ N_2\ !\ \ldots\ N_\varrho\ !} \tag{19}$$

where N is the total number of individuals and N_ϱ is the number of speakers of the ϱth language.

D_B may be interpreted, similarly to equation (2), as the language diversity per individual. Unlike Shannon's index, D_B increases as a function of N. This should not be viewed as a drawback, however, since it is not unreasonable to expect large populations to be more diverse than small ones.

The use of logarithms ensures that D_B has the property of additivity, (3) (page 4) in addition to properties (1) and (2). In the case of very large values of N Pielou (1969) suggests the use of Stirling's approximation to the factorial

$$\ln N! \simeq N(\ln N - 1) + \tfrac{1}{2} \ln 2\pi N.$$

Thus, should it be required to take into account the hierarchical nature of language classification in a highly diverse unit such as India, for which census data were available,

one would initially define, diversity of the populations in terms of their affiliation to language groups as

$$(20) \qquad D_G = \frac{1}{N} \log \frac{N\ !}{N_1\ !\ N_2\ !\ldots N_i\ !\ldots\ N_g\ !}$$

where N_1 might be the number of individuals in the Dravidian group, N_2 those in the Indo-European group and so-on.

Language diversity *within* the ith language group would then be defined as

$$(21) \qquad D_{(L)_i} = \frac{1}{N_i} \log \frac{N_i\ !}{N_{i_1}\ !\ N_{i_2}\ !\ldots N_{i_j}\ !\ldots N_{i_{\ell i}}\ !}$$

where there are ℓi languages in the ith group, N_i is the number of individuals in language group i, and N_{i_j} is the number of users of language j. Total diversity would be given by

$$(22) \qquad D_{tot.} = \frac{1}{N} \log \frac{N\ !}{\prod_{j=1}^{\ell 1} N1j\ !\ \prod_{j=1}^{\ell 2} N2j\ !\ldots\ \prod_{j=1}^{\ell g} Ngj\ !}$$

multiplying (22) by $\dfrac{N_1\ !\ldots N_g\ !}{N_1\ !\ldots N_g\ !}$

we obtain $$D_{tot.} = \frac{1}{N} \log \left(\frac{N\ !}{N_1\ !\ldots N_g\ !} \cdot \frac{N_1\ !\ldots N_g\ !}{\prod_{j=1}^{\ell 1} N1j\ !\ldots \prod_{j=1}^{\ell g} Ngj\ !} \right)$$

$$D_{tot.} = D_G + \frac{1}{N} \sum_{i=1}^{g} \log \frac{N_i\ !}{\prod_{j=1}^{\ell i} Nij\ !}$$

multiplying by $\frac{Ni}{Ni}$ we obtain

$$D_{tot.} = D_G + \sum_{i=1}^{g} \frac{N_i}{N} D(L)_i \qquad (23)$$

Equation (23) may be extended to take in three (or more) levels of language classification such as family, group and language as follows

$$D = D_F + \sum_{i=1}^{f} \frac{N_i}{N} D_{(G)_i} + \sum_{i=1}^{f} \sum_{j=1}^{gi} \frac{N_{ij}}{N} D_{(L)_{ij}} \qquad (24)$$

In order that the non-mathematically inclined reader may find his way through this welter of subscripts, a visual example of the application of (24) to a hypothetical, threefold language classification is shown in figure 1.

While the ability to allow for the hierarchical nature of language classification would prove useful at a continental or world scale, equation (19) will probably prove adequate in most instances. The degree of evenness of the distribution of fully censused individuals among languages may be evaluated by calculating D_B / D_{max}. (Cf. equation 6). Pielou (1975, 1977) shows D_{Bmax} to be

$$D_{Bmax} = \frac{1}{N} \log \frac{N!}{(X!)^{\ell - r} (Y!)^r} \qquad (25)$$

where $X = [N/\ell]$ or the integer part of the total number of individuals divided by the number of languages, and $Y = X + 1$, so that $N = (\ell - r) X + r Y$. For the purposes of census language data analysis this expression should of course be simplified, entailing no noticeable loss of precision, to

$$D_{Bmax} = \frac{1}{N} \log \frac{N!}{(X!)^{\ell}} \qquad (26)$$

It will be recalled that the equations based on Brillouin's index are appropriate to the measurement of the diversity of fully censused populations and that there is therefore no need to calculate their standard error.

Some Other Approaches to Diversity Measurement

Pielou (1969) quotes a geometrical interpretation of the concept of diversity introduced by McIntosh (1967). McIntosh suggested that a population consisting of N individuals and ℓ discrete categories with N_i individuals in the ith category may be interpreted as a point in an ℓ-dimensional space with coordinates $(N_1, N_2, N_i, \ldots N_\ell)$. The distance

of this point from the origin, by Pythagoras' theorem, is

(27) $$H = \sqrt{\Sigma\ N_i^2}$$

The greater the number of categories (ℓ) the smaller is the distance. H may therefore be interpreted as a measure of the language homogeneity of the population. $H_{max.}$ will be attained when all individuals are assigned to a single category ($H = N$). $H_{min.}$ will be attained when every individual speakes a different language ($N = \sqrt{N}$). As with the Greenberg/Simpson index this measure would describe language homogeneity rather than language diversity. It is appropriate therefore to take

(28) $$D = N - H$$

as a measure of diversity. McIntosh further proposed

(29) $$\frac{D}{D_{max}} = \frac{N - H}{N - \sqrt{N}}$$

as a measure of diversity which is independent of the population size, N.

McIntosh also developed a measure of evenness of distribution of individuals among categories. Assuming that the number of categories (languages) may be divided into the total population exactly, resulting in $\frac{N}{\ell}$ individuals using each language then

(30) $$\text{Hmin for given } N, \ell = \left(\Sigma(\frac{N}{\ell})^2\right)^{\frac{1}{2}} = \frac{N}{\sqrt{\ell}}$$

The N - complement of Hmin for given N,ℓ is thus

(31) $$D_{max} \text{ for given } N, \ell = N - N/\sqrt{\ell}$$

The degree of evenness of a given distribution may therefore by measured by

(32) $$\frac{D}{D_{max}|N,\ell} = \frac{N - H}{N - N/\sqrt{\ell}}$$

McIntosh's index (27) is somewhat reminiscent of Weinreich's suggested D measure (1957)

(33) $$D = 1 - \sqrt{\Sigma pi^2}$$

although Weinreich advocated the use of proportions rather than of numbers of users. Weinreich's D is of course also a transformation of Greenberg's A-index; its superiority to the latter author's measure is not readily apparent, however, and it has not been employed elsewhere to the present writer's knowledge.

As a final example of the more promising methods of measuring language diversity, it is perhaps worthwhile pointing out that, in some cases, the standard deviation of the distribution of numbers of users of each language could constitute a quantitative indication of the diversity of a set of languages and their users. It would have to be interpreted with care, however, because of the weighting of extreme deviations from the mean which would be likely in this context. Furthermore, it could only be used for the comparison of societal units featuring roughly similar numbers of languages. It might nevertheless be considered for the measurement of change in diversity within a given unit over time, and has the advantage of being widely known.

Table 1 summarises the definitions and applicability of the various indices reviewed above.

The foregoing section has reviewed some of the more promising ways of measuring language diversity which are currently available to the investigator. The following pages discuss some of the pitfalls which may be encountered in seeking to apply such measures.

Problems in the Application of a Diversity Index

1. *Classification of individuals*

One difficulty in the application of a diversity index is the unambiguous assignment of individuals to a discrete language class. A prerequisite for this, as stated on page (2), is a clear statement of the subset of linguistic activities for which a diversity measurement is required, since different languages may be used for different activities. In the case of individuals using more then one language for a given activity, additional multilingual categories could be created (Greenberg, 1956). A further requirement is that languages may in fact be subjected to a discrete classification, a situation which has recently been attained, at least, on a genetic basis (C.F. and F.M. Voegelin, 1977).

2. *The modifiable unit area problem*

More intractable difficulties are raised by the definition of the areal unit for which a diversity measurement is to be made. Obviously, the measurement of the language diversity of a set of human beings requires a certain minimum level of aggregation. But, as a

TABLE 1

Index	*Definition*	*Type of data*
Shannon	$- \Sigma \frac{N_i}{N} \log \frac{N_i}{N}$	Sample data
Brillouin	$\frac{1}{N} \log \frac{N\ !}{N_1\ !\ :\ N_i\ !\ \ldots\ N_\ell\ !}$	Census data
Simpson/Greenberg	$1 - \Sigma \left(\frac{N_i}{N}\right)^2$	Sample or census data
McIntosh	$N - \sqrt{\Sigma N_i^2}$	Census data
Stand. Dev.	$\sqrt{\frac{\Sigma (N_i - N)^2}{\ell - 1}}$	Sample data (or census data with denominator = ℓ).
Weinreich	$1 \sqrt{\Sigma \left(\frac{N_i}{N}\right)^2}$	Census data

N_i = Number of users of ith language, N = total population, ℓ = number of languages.

perusal of any modern textbook in theoretical geography will reveal (e.g. Yeats, 1974, Taylor 1977, Harvey, 1969) the parameters of a spatial distribution will change according to the level of aggregation at which it is examined, often quite substantially (cf. Lieberson and O'Connor, 1975a). This, of course, is the areal aspect of the celebrated ecological fallacy (Robinson, 1950, Duncan *et al.* 1961, Scheuch, 1966). There is not space to go into the general problem in detail here, the reader will find a thorough discussion in Duncan *et al.* (1961). Scheuch sums the matter up succintly when he states :

> In the logic of inquiry it does not make any difference whether the basis for grouping individual units is a territory or some other criterion; what is essential is the effect that this criterion has on the control over the internal variability of units. Thus the general issue underlying the discussion of the ecological fallacy is really the relation of the criterion, according to which units are grouped, to the type of inference intended when using the results of aggregated units.[5]

[5] Scheuch (1966), p. 154.

Language data are usually collected on a local areal basis (census, tracts, counties, etc.) and subsequently aggregated to a greater or lesser degree prior to final presentation. The question is, at what level should the investigator work ? In the case of repeated diversity measurements within a single observational unit over time no great difficulty should be encountered, providing the unit is a meaningful one in terms of language diversity. Weinreich suggests that « ideally, boundaries should reflect actual communication patterns »[6]. Such a unit will not necessarily correspond to the administratively determined boundaries of most government data collection areas, of course. In the case of inter-unit comparisons, however, care must be exercised as such a procedure may mean that the researcher is comparing results which were obtained at radically different levels of aggregation. Greenberg (1956) pointed this out in his pioneering paper, but his advice has not always been heeded, particularly insofar as crosspolity comparisons are concerned. Is it meaningful, one may ask, to compare the language diversity of Eire (pop. < 3 mill.) with that of the U.S.S.R. (pop. > 220 mill.) ?

The modifiable unit area problem is particularly serious when one is interested in relating a measure of diversity to other variables. Lieberson and Hansen (1974), for example, examine the relationship between diversity (Greenberg's A-index) and urbanisation in the case of the U.S.S.R. at several points in time. No noticeable correlation emerges, and it is concluded that the two variables are not related. This is no doubt perfectly true, at the scale of the U.S.S.R. It must be borne in mind, however, that the linguistically diverse regions of the U.S.S.R. have not generally coincided with those characterised by high degrees of urbanisation. Percentages of city dwellers at any one time may have tended to reflect largely the situation in European Russia, while language diversity readings may have been swollen by the linguistic situation in other areas such as Soviet Central Asia. To some extent, therefore, the correlated measurements may not have referred unambiguously to the same groups of individuals, and, consequently, that which is valid at the continental scale of the U.S.S.R. may not hold at the regional scale of, say, Kazakhstan. It is interesting to note that, at the finer, regional scale, Lieberson, Dalto and Johnston (1975b) do report a noticeable, positive diversity/urbanisation correlation within the U.S.S.R.

It is even more difficult to assess the validity of cross-national (i.e. spatial) correlations where the levels of aggregation of the units of observation vary wildly among themselves (e.g. Liechtenstein and India, Luxembourg and Canada, etc.). A number of writers including the present author (Brougham, 1969) have viewed the scale phenomenon as a feature to be studied in its own right rather than as a problem to be eliminated. The concept of language diversity after all, presupposes data aggregation, and its relation to other variables implies ecological rather than individual correlations. It is therefore advisable, wherever possible, to measure diversity, and hypothesised explanatory variables, at several levels of aggregation in order to discover at which scale, if any, variation, order and relationships exist. Every effort should be made to avoid spatial correlations of language diversity measurements with other variables when units of observation are at totally incompatible levels of data aggregation.

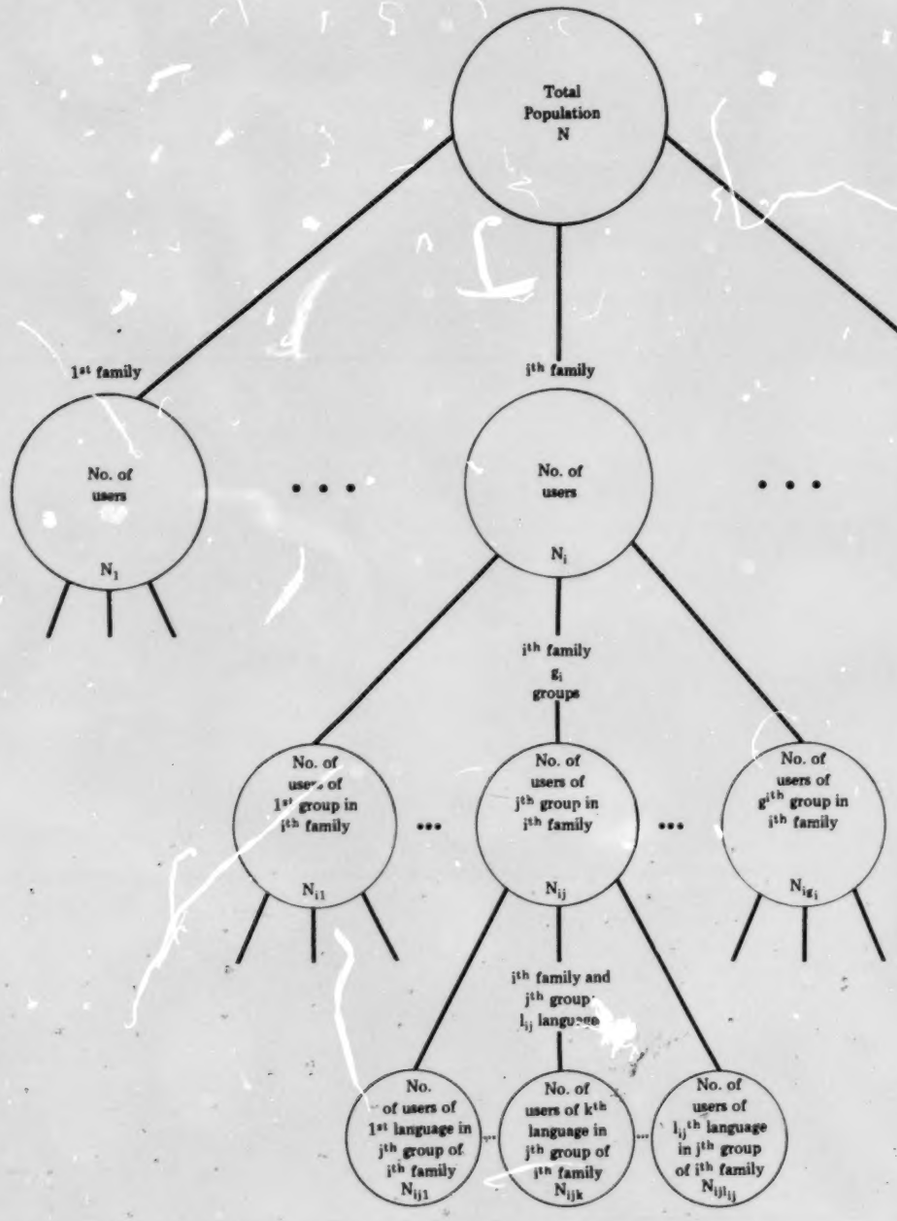
[6] Weinreich (1957), p. 227.

CONCLUSION

The preceding pages have presented the properties which one might reasonably expect a language diversity index to exhibit, and have reviewed some of the indices which are presently available for the measurement of this phenomenon, using either sample or census data. It was further suggested that one of the main problems in the measurement of diversity, and its subsequent correlation with other variables, is the definition of the observational units. Ways of attenuating this difficulty are currently being investigated at the CIRB and will be presented shortly. It is also planned to apply the more promising of the diversity indices reviewed here to the data accumulated at the Centre by Kloss & McConnell (1974, 1978, 1979) in their study of the linguistic composition of the nations of the world, thereby providing a picture of language diversity across the globe at various scales. It is also hoped that other workers will find these indices of value in terms of the evaluation and explanation of the phenomenon of language diversity.

FIGURE 1:

HYPOTHETICAL EXAMPLE OF DIVERSITY MEASUREMENT FOR A THREEFOLD LANGUAGE CLASSIFICATION USING BRILLOUIN'S INDEX



$$D_p = \frac{1}{N} \log \frac{N!}{N_1! \cdots N_i! \cdots N_f!}$$

$$D_{(G)i} = \frac{1}{N_i} \log \frac{N_i!}{N_{i1}! \cdots N_{ij}! \cdots N_{ig_i}!}$$

$$D_{(L)ij} = \frac{1}{N_{ij}} \log \frac{N_{ij}!}{N_{ij1}! \cdots N_{ijk}! \cdots N_{ijl_{ij}}!}$$

$$D_{tot.} = D_p + \sum_{i=1}^{f} \frac{N_i}{N} D_{(G)i} + \sum_{i=1}^{f} \sum_{j=1}^{g_i} \frac{N_{ij}}{N} D_{(L)ij}$$

REFERENCES

BANKS, Arthur S., and TEXTOR, Robert B., (1963) *A Cross-Polity Survey*. Cambridge, M.I.T.

BOWMAN, K.O., HUTCHESON, K., ODUM, E.P. and SHENTON, L.R., (1971) « Comments on the Distribution of Indices of Diversity », *Proceedings of the International Symposium on Statistical Ecology*. Pennsylvania State University Press.

BRILLOUIN, L., (1962) *Science and Information Theory*. New York, Academic, 2nd ed.

BROUGHAM, James (1969) « Impact of Crowding upon the Quality of the Recreational Experience ». Unpublished M.A. thesis, University of Western Ontario.

CRIPER, Clive, and LADEFOGED, Peter, (1971) « Linguistic Complexity in Uganda », in *Language Use and Social Change*. Oxford University Press.

DEUTSCH, Karl W., (1966) *Nationalism and Social Communications : an Enquiry into the Foundations of Nationality*. Cambridge, M.I.T., 2nd ed.

DUNCAN, Otis, DUDLEY, Cuzzort, RAY, P., and DUNCAN, Beverly (1961) *Statistical Geography*. Connecticut, Greenwood Press.

FISHMAN, Joshua A., (1968a) « Nationality-nationalism and Nation-nationism », in *Fishman et al.*, 39-51.

—— (1968b) « Some Contrasts between Linguistically Homogeneous and Linguistically Heterogeneous Politics », in *Fishman et al.*, 53-68.

FISHMAN, Joshua A., FERGUSON, Charles A., and DAS GUPTA, Jyotirinda (eds). (1968) *Language Problems of Developing Nations*. New York, Wiley.

FISHMAN, Joshua A., COOPER, Robert L., and ROSENBAUM, Yehudit, (1977) « English the World Over : A Factor in the Creation of Bilingualism Today », in *Bilingualism : Psychological, Social & Educational Implications*. New York, Academic Press Inc.

GREENBERG, Joseph H., (1956) « The Measurement of Linguistic Diversity », *Language* 32 (1), 109-115.

GUMPERZ, John J., (1962) « Language Problems in the Rural Development of North India », in *Rice*, 79-90.

HARVEY, David, (1969) *Explanation in Geography*. Toronto, Macmillan.

HILL, M.O., (1973) « Diversity and Evenness : A Unifying Notation and its Consequences », *Ecology* 54, 427-432.

HUTCHESON, Kesmit, (1970) « A Test for Comparing Diversities Based on the Shannon Formula », *Journal of Theoretical Biology*, 29, 151-154.

KLOSS, H., and McCONNEL, G.D., (1974, 1978, 1979) *Linguistic Composition of the Nations of the World*. Québec, CIRB, Presses de l'Université Laval.

LIEBERSON, Stanley, (1968) « An Extension of Greenberg's Linguistic Diversity Measures », in *Readings in the Sociology of Language*. Ed. Joshua Fishman, The Hague, Mouton.

LIEBERSON, Stanley, and HANSEN, Lynn H., (1974) « National Development, Mother Tongue Diversity, and the Comparative Study of Nations », *American Sociological Review*, Vol. 39 (August), 523-541.

LIEBERSON, Stanley, and O'CONNOR, James F.C., (1975a) « Language Diversity in a Nation and its Regions », in *Multilingual Political Systems : Problems and Solutions*, eds. Savard, Jean-Guy & Vigneault, Richard, Québec, Presses de l'Université Laval.

LIEBERSON, Stanley, DATTO, Guy and JOHNSTON, Mary Ellen, (1975b) « The Course of Mother Tongue Diversity in Nations », *American Sociological Review*, Vol. 81 (July), 34-61.

McINTOSH, R.P., (1967) « An Index of Diversity & the Relation of Certain Concepts to Diversity », *Ecology* 48, 302-404.

PIELOU, E.C., (1966) « The Measurement of Diversity in Different Types of Biolgical Collections », *Journal of Theoretical Biology*, 13, 131-144.

—— (1967) « The Use of Information Theory in the Study of the Diversity of Biological Populations », *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 4, 163-177.

—— (1969) *An Introduction to Mathematical Ecology*. New York, Wiley.

—— (1975) *Ecological Diversity*. New York, Wiley.

—— (1977) *Mathematical Ecology*. New York, Wiley.

POOL, Jonathan, (1969) « National Development and Language Diversity », *La Monda Linguo-Problemo* 1 (3), 140-156.

POOLE, Robert W., (1974) *An Introduction to Quantitative Ecology*. McGraw Hill.

RENYI, Alfred, (1961) « On Measures of Entropy and Information », *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, I, 547-561.

RICE, Frank A., (1962) *Study of the Role of Second Languages*. Washington, Center for Applied Linguistics.

ROBINSON, W.S., (1950) « Ecological Correlations and the Behavior of Individuals », *American Sociological Review*, 15, 351-357.

SADLER, Victor, (1962) « A Measure of 'Internationality' », *International Associations*, (7), 480-481.

SCHEUCH, Erwin K., (1966) « Cross National Comparisons & Using Aggregate Data : Some Substantive and Methodological Problems », in *Comparing Nations : The Use of Quantitative Data in Cross-National Research*. Eds. Merritt, Richard L. & Rolekan, Strein. Yale University Press.

SHANNON, C.E., and WEAVER, W., (1949) *The Mathematical Theory of Communication*. Urbana, University of Illinois Press.

SIMEON, Richard, (1962) « The Scope and Development of Government Policy », in *Linguistic & Cultural Diversity*, Canadian Commission for Unesco, Ottawa, 14-25.

SIMPSON, E.H., (1949) « Measurement of Diversity », *Nature* 163.

TAYLOR, P.J., (1977) *Quantitative Methods in Geography*, Boston, Houghton Mifflin.

VOEGELIN, C.F., and VOEGELIN, F.M., (1977) *Classification and Index of the World's Languages*, Elsevier.

WEAVER, W., (1948) « Probability, Rarity, Interest and Surprise », *The Scientific Monthly* 67, 390-392.

WEINREICH, Uriel, (1957) « Functional Aspects of Indian Bilingualism », *Word*, Vol. 13 (2), 203-233.

YEATES, Maurice, (1974) *An Introduction to Quantitative Analysis in Human Geography*, Toronto, McGraw-Hill.

Dans la même série:

B-1 *L'utilisation de l'ordinateur en lexicométrie.*
Savard, Jean-Guy

B-2 *L'ordinateur et l'analyse grammaticale.*
Mepham, Michael S.

B-3 *Concept Categories as Measures of Culture Distance.*
Mackey, William F.

B-4 *L'université bilingue.*
Verdoodt, Albert

B-5 *La rentabilité des mini-langues.*
Mackey, William F.

B-6 *The Computer in Automated Language Teaching.*
Mackey, William F.

B-7 *The Three-Fold Objective of the Language Reform in Mainland China in the Last Two Decades.*
Chiu, Rosaline Kwan-wai

B-8 *Un test télévisé.*
Savard, Jean-Guy

B-9 *Sociolinguistic History, Sociolinguistic Geography and Bilingualism.*
Afendras, Evangelos A.

B-10 *Mathematical Models for Balkan Phonological Convergence.*
Afendras, Evangelos A.

B-11 *Stability of a Bilingual Situation and Arumanian Bilingualism.*
Afendras, Evangelos A.

B-12 *More on Informational Entropy, Redundancy and Sound Change.*
Afendras, Evangelos A. & Tzannes, Nicolaos S.

B-13 *Relations entre anglophones et francophones dans les syndicats québécois.*
Verdoodt, Albert

B-14 *Multilingual Communication in Nigeria.*
Iso, Asi Otu & Afendras, Evangelos A.

B-15 *The Language Factor in Maori Schooling.*
Richards, Jack C.

B-16 *Diffusion Processes in Language: prediction and planning.*
Afendras, Evangelos A.

B-17 *A Non-Contrastive Approach to Error Analysis.*
Richards, Jack C.

B-18 *Research Possibilities on Group Bilingualism: a report.*
Kloss, Heinz & Verdoodt, Albert

B-19 *Interference, Integration and the Synchronic Fallacy.*
Mackey, William F.

B-20 *A Psycholinguistic Measure of Vocabulary Selection.*
Richards, Jack C.

B-21 *A Pilot Study on the Ability of Young Children and Adults to Identify and Reproduce Novel Speech Sounds.*
Afendras, Evangelos A., Yeni-Komshian, G. & Zubin, David A.

B-22 *Can One Measure a Sprachbund? A Calculus of Phonemic Distribution for Language Contact.*
Afendras, Evangelos A.

B-23 *Stochastic Processes for Diachronic Linguistics.*
Afendras, Evangelos A. & Tzannes, Nicolaos S.

B-24 *Structures ethniques et linguistiques au Burundi, pays 'unimodal' typique.*
Verdoodt, Albert

B-25 *Error Analysis and Second Language Strategies.*
Richards, Jack C.

B-26 *Graduate Education in Foreign Language Teaching.*
Mackey, William F.

B-27 *La question scolaire en Alsace: statut confessionnel et bilinguisme.*
Kauffmann, Jean

B-28 *Polychronometry: the study of time variables in behavior.*
Mackey, William F.

B-29 *Diglossie au Québec: limites et tendances actuelles.*
Chantefort, Pierre

B-30 *Literary Biculturalism and the Thought-Language-Culture Relation.*
Mackey, William F.

B-32 *La distance interlinguistique.*
Mackey, William F.

B-33 *Options politiques fondamentales de l'état plurilingue.*
Plourde, Gaston

B-34 *Social Factors, Interlanguage and Language Learning.*
Richards, Jack C.

B-35 *Analyse des erreurs et grammaire générative: la syntaxe de l'interrogation en français.*
Py, Bernard

B-36 *Anglicization in Quebec City.*
Edwards, Vivien

B-37 *La lexicométrie allemande: 1898-1970.*
Njock, Pierre-Emmanuel

B-39 *Individualisation de l'enseignement et progrès continu à l'élémentaire. Application à l'anglais, langue seconde.*
Bégin, Y., Masson, J.P., Beaudry, R. & Paquet, D. (INRS-Education)

B-41 *Une communauté allemande en Argentine: Eldorado.*
Micolis, Marisa

B-42 *Three Concepts for Geolinguistics.*
Mackey, William F.

B-43 *Some Formal Models for the Sociology of Language: diffusion, prediction and planning of change.*
Afendras, Evangelos A.

B-45 *Le projet de restructuration scolaire de l'île de Montréal et la question linguistique au Québec.*
Duval, Lise & Tremblay, Jean-Pierre; recherche dirigée par Léon Dion avec la collaboration de Micheline de Sève.

B-46 *L'écologie éducationnelle du bilinguisme.*
Mackey, William F.

B-47 *La situation du français comme langue d'usage au Québec.*
Gendron, Jean-Denis

B-48 *Network Concepts in the Sociology of Language.*
Afendras, Evangelos A.

B-49 *Attitude linguistique des adolescents francophones du Canada.*
Gagnon, Marc

B-50 *Vers une technique d'analyse de l'enseignement de l'expression orale.*
Huot-Tremblay, Diane

B-51 *A Demographic Profile of the English Population of Quebec 1921-1971.*
Caldwell, Gary

B-52 *Language in Education and Society in Nigeria: a comparative bibliography and research guide.*
Brann, C.M.B.

B-53 *Eléments de correction phonétique du français.*
LeBel, Jean-Guy

B-54 *Langue, dialecte et diglossie littéraire.*
Mackey, William F.

B-55 *Rapport de synthèse de l'élaboration du test d'anglais langue seconde.*
GREDIL (Groupe de recherche et d'étude en didactique des langues)

B-56 *Relations interethniques et problèmes d'acculturation.*
Abou, Sélim

B-57 *Etude socio-linguistique sur l'intégration de l'immigrant allemand au milieu québécois.*
Hardt-Dhatt, Karin

B-58 *La culture politique du Mouvement Québec Français.*
Turcotte, Denis

B-59 *Aspects sociolinguistiques du bilinguisme canadien.*
Saint-Jacques, Bernard

B-60 *Cooperation and Conflict in Dual Societies: a comparison of French-Canadian and Afrikaner nationalism.*
Novek, Joël

B-61 *Le Zaïre: deuxième pays francophone du monde?*
Faïk, Sully; Pierre, Max; N'Tita, Nyembwe & N'Sial, Sesep

B-62 *7e Colloque 1976 — Actes / 7th Symposium 1976 — Proceedings.*
Association canadienne de linguistique appliquée /
Canadian Association of Applied Linguistics

B-63 *Les dispositions juridico-constitutionnelles de 147 Etats en matière de politique linguistique.*
Turi, Giuseppe

B-64 *Contribution à l'étude du problème de la difficulté en langue étrangère.*
Ragusich, Nicolas-Christian

B-65 *Linguistic Tensions in Canadian and Belgian Labor Unions.*
Verdoodt, Albert

B-66 *Contribution à l'étude de la nouvelle immigration libanaise au Québec.*
Abou, Sélim

B-67 *L'incidence de l'âge dans l'apprentissage d'une langue seconde.*
Daigle, Monique

B-68 *The Contextual Revolt in Language Teaching.*
Mackey, William F.

B-69 *La langue française en Afrique occidentale francophone.*
Kwofie, Emmanuel N.

B-70 *Motivational Characteristics of Francophones Learning English.*
Clément, Richard

B-71 *Schedules for Language Background, Behavior and Policy Profiles.*
Mackey, William F.

B-72 *Difficultés phonétiques de l'acquisition du français, langue seconde.*
Huot, France

B-73 *Multilinguisme et éducation au Nigéria.*
Brann, C.M.B.

B-74 *Les systèmes approximatifs et l'enseignement des langues secondes.*
High Locastro, Virginia

B-75 *Le bilinguisme canadien: bibliographie analytique et guide du chercheur.*
Mackey, William F.

B-76 *Un siècle de colloques sur la didactique des langues.*
Mackey, William F.

B-77 *L'irrédentisme linguistique: une enquête témoin.*
Mackey, William F.

B-78 *Babel: perspectives for Nigeria.*
Simpson, Ekundayo

B-79 *Samuel Beckett: traducteur de lui-même.*
Simpson, Ekundayo

B-80 *8e Colloque 1977 — Actes / 8th Symposium 1977 — Proceedings*.
Association canadienne de linguistique appliquée /
Canadian Association of Applied Linguistics

B-81 *Language Survey for Nigeria*.
Osaji, Bede

B-82 *L'univers familier de l'enfant africain*.
Njock, Pierre-Emmanuel

B-83 *The Social Psychology of Inter-ethnic Contact and Cross-cultural Communication: An Annotated Bibliography*.
Desrochers, Alain & Clément, Richard

B-84 *Géographie du français et de la francité en Louisiane*.
Breton, Roland J.-L.

B-85 *Etude morphosyntaxique du parler acadien de la Baie Sainte-Marie, Nouvelle-Ecosse (Canada)*.
Gesner, B. Edward

B-86 *Multinational Schools as Language Learning Media*.
Mackey, William F.

B-87 *Translating in the Nigerian Mass Media: A Sociolinguistic Study*.
Simpson, Ekundayo

B-88 *Identité culturelle et francophonie dans les Amériques (III)*.
Baudot, Alain, Jaubert, Jean-Claude & Sabourin, Ronald

B-89 *Les banques de terminologie bilingues et multilingues: Etat de la question*.
Rondeau, Guy

B-90 *Differences in Earnings by Language Groups in Quebec, 1970: An Economic Analysis*.
Vaillancourt, François

B-91 *The Role of France, Quebec and Belgium in the Revival of French in Louisiana Schools*.
Gold, Gerald L.

B-92 *L'éducation des enfants de travailleurs migrants en Europe occidentale (Bibliographie sélective)*.
Rosseel, Eddy

B-93 *La distance interlinguistique lexicale*
Huot, Jean-Claude

B-94 *Le français parlé: analyse des attitudes des adolescents de la ville de Québec selon les classes sociales*.
Noël, Dany (Danièle)

B-95 *Bilingualism and Linguistic Segregation in the Schools of Brussels*.
Elizabeth Sherman Swing

B-96 *Apprentissage dans des contextes bilingues*.
Rodrigue Landry

B-97 *Exogamie et anglicisation dans les régions de Montréal, Hull, Ottawa et Sudbury*.
Castonguay, Charles

AUTRES PUBLICATIONS DU C.I.R.B.

Série A — Etudes/Studies (Presses de l'Université Laval)

*A-1 SAVARD, Jean-Guy et RICHARDS, Jack C. *Les indices d'utilité du vocabulaire fondamental français.* Québec, 1970, 172 p.

A-2 KLOSS, Heinz. *Les droits linguistiques des Franco-Américains aux Etats-Unis.* Québec, 1971, 84 p.

A-3 FALCH, Jean. *Contribution à l'étude du statut des langues en Europe.* Québec, 1973, 284 p.

A-4 DORION, Henri & MORISSONNEAU, Christian (colligés et présentés/editors). *Les noms de lieux et le contact des langues / Place Names and Language Contact.* Québec, 1972, 374 p.

A-5 LAFORGE, Lorne. *La sélection en didactique analytique.* Québec, 1972, 383 p.

A-6 TOURET, Bernard. *L'aménagement constitutionnel des Etats de peuplement composite.* Québec, 1973, 260 p.

A-7 MEPHAM, Michael S. *Computation in Language Text Analysis.* Québec, 1973, 234 p.

A-8 CAPPON, Paul. *Conflit entre les Néo-Canadiens et les francophones de Montréal.* Québec, 1974, 288 p.

A-9 SAVARD, Jean-Guy & VIGNEAULT, Richard (présentation/presentation). *Les états multilingues: problèmes et solutions / Multilingual Political Systems: problems and solutions.* Textes de la Table Ronde de 1972/Papers of the Round Table in 1972. Québec, 1975, 591 p.

A-10 BRETON, Roland J.-L. *Atlas géographique des langues et des ethnies de l'Inde et du subcontinent.* Québec, 1976, 648 p.

A-11 SNYDER, Emile & VALDMAN, Albert (présentation). *Identité culturelle et francophonie dans les Amériques.* Québec, 1976, 290 p.

A-12 DARBELNET, Jean. *Le français en contact avec l'anglais en Amérique du Nord.* Québec, 1976, 146 p.

A-13 MALLEA, John R. (compiled and edited). *Quebec's Language Policies: background and response.* Québec, 1977, 309 p.

A-14 DORAIS, Louis-Jacques. *Lexique analytique du vocabulaire inuit moderne au Québec-Labrador.* Québec, 1978, 136 p.

A-15 CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME / INTERNATIONAL CENTER FOR RESEARCH ON BILINGUALISM. *Minorités linguistiques et interventions: Essai de typologie / Linguistic Minorities and Interventions: Towards a Typology.* Compte rendu du Colloque sur les minorités linguistiques tenu à l'Université Laval du 15 au 18 avril 1977 / Proceedings of the Symposium on Linguistic Minorities held at Laval University from April 15th to April 18th 1977. Québec, 1978, 318 p.

*Epuisé / Out of print

A-100 EQUIPE DE PROFESSEURS DE L'UNIVERSITE LAVAL. *Test Laval: formule A, classement en français langue seconde*. Québec, 1971, Copie échantillon/Sample copy.

A-101 EQUIPE DE PROFESSEURS DE L'UNIVERSITE LAVAL. *Test Laval: formule B, formule C, classement en français langue seconde*. Québec, 1976, Copie échantillon/Sample copy.

Série C – Publications extérieures/Outside publications

C-1 SAVARD, Jean-Guy. *La valence lexicale*. Paris, Didier, 1970, 236 p.

C-2 MACKEY, William F. *Le bilinguisme phénomène mondial / Bilingualism as a World Problem*. Montréal, Harvest House, 1967, 119 p.

C-3 MACKEY, William F., SAVARD, Jean-Guy & ARDOUIN, Pierre. *Le vocabulaire disponible du français*. Montréal, Didier Canada, 1971, 2 volumes, 900 p.

C-4 STERN, H.H. (rédacteur). *L'enseignement des langues et l'écolier. Rapport d'un colloque international*. (Traduit au CIRB sous la direction de William F. Mackey). Hambourg, Institut de l'UNESCO pour l'éducation, 1971, 254 p.

C-5 KLOSS, Heinz. *Laws and Legal Documents Relating to Problems of Bilingual Education in the United States*. Washington, D.C., Center for Applied Linguistics, 1971, 92 p.

C-6 MACKEY, William F. *Principes de didactique analytique*. (Révisé et traduit par Lorne Laforge). Paris, Didier, 1972, 713 p.

C-7 MACKEY, William F. & VERDOODT, Albert (editors). *The Multinational Society*. Rowley (Mass.), Newbury House, 1975, 388 p.

C-8 GIORDAN, Henri & RICARD, Alain (sous la direction). *Diglossie et littérature*. Bordeaux-Talence, Maison des sciences de l'homme, 1976, 184 p.

C-9 MACKEY, William F. *Bilinguisme et contact des langues*. Paris, Klincksieck, 1976, 539 p.

C-10 MACKEY, William F., ORNSTEIN, Jacob & al. *The Bilingual Education Movement: essays on its progress*. El Paso, Texas Western Press, 1977, 153 p.

C-11 MACKEY, William F., & ORNSTEIN, Jacob (editors). *Sociolinguistic Studies in Language Contact*. The Hague, Mouton, 1979, 460 p.

Collection *Studies in Bilingual Education* (Newbury House, Rowley, Mass.)
W.F. Mackey – General Editor

C-100 MACKEY, William F. *Bilingual Education in a Binational School. a study of equal language maintenance through free alternation*. 1972, 185 p.

C-101 SPOLSKY, Bernard (editor). *The Language Education of Minority Children: selected readings*. 1972, 200 p.

C-102 LAMBERT, Wallace E. & TUCKER, G. Richard. *Bilingual Education of Children: the St. Lambert experiment*. 1972, 248 p.

C-103 COHEN, Andrew D. *A Sociolinguistic Approach to Bilingual Education: Experiments in the American Southwest*. 1975, 352 p.

C-104 GAARDER, A. Bruce. *Bilingual Schooling and the Survival of Spanish in the United States*. 1977, 238 p.

C-105 KLOSS, Heinz. *The American Bilingual Tradition*. 1977, 347 p.

C-106 MACKEY, William F. & ANDERSSON, Theodore. *Bilingualism in Early Childhood*. 1977, 443 p.

C-107 MACKEY, William F. & BEEBE, Von-Nieda. *Bilingual Schools for a Bicultural Community*. 1977, 223 p.

Série E – Inventaires/Inventories (Presses de l'Université Laval)

E-1 KLOSS, Heinz & McCONNELL, Grant D. (rédacteurs/editors). *Composition linguistique des nations du monde.* Vol. 1: *L'Asie du Sud: secteurs central et occidental / Linguistic Composition of the Nations of the World.* Vol. 1: *Central and Western South Asia.* Québec, 1974, 408 p.

E-2 KLOSS, Heinz & McCONNELL, Grant D. (rédacteurs/editors). *Composition linguistique des nations du monde.* Vol. 2: *L'Amérique du Nord / Linguistic Composition of the Nations of the World.* Vol. 2: *North America.* Québec, 1978, 893 p.

E-3 KLOSS, Heinz & McCONNELL, Grant D. (rédacteurs/editors). *Composition linguistique des nations du monde.* Vol. 3: *L'Amérique centrale et l'Amérique du Sud / Linguistic Composition of the Nations of the World.* Vol. 3: *Central and South America.* Québec, 1979, 564 p.

E-10 KLOSS, Heinz & McCONNELL, Grant D. (rédacteurs/editors). *Les langues écrites du monde: relevé du degré et des modes d'utilisation.* Vol. 1 *Les Amériques / The Written Languages of the World: a survey of the degree and modes of use.* Vol. 1: *The Americas.* Québec, 1978, 633 p.

(à paraître/forthcoming)

E-4 KLOSS, Heinz & McCONNELL, Grant D. (rédacteurs/éditors). *Composition linguistique des nations du monde.* Vol. 4: *L'océanie / Linguistic Composition of the Nations of the World.* Vol. 4: *Oceania.* Québec.

Série F – Bibliographies (Presses de l'Université Laval)

F-1 SAVARD, Jean-Guy. *Bibliographie analytique de tests de langue / Analytical Bibliography of Language Tests.* Québec, 2e éd., 1977, 570 p.

*F-2 CHIU, Rosaline Kwan-wai. *Language Contact and Language Planning in China (1900-1967). A Selected Bibliography.* Québec, 1970, 276 p.

F-3 MACKEY, William F. (rédacteur/editor). *Bibliographie internationale sur le bilinguisme / International Bibliography on Bilingualism.* Québec, 1972, 757 p.

F-4 AFENDRAS, Evangelos A. & PIANAROSA, Albertina. *Bibliographie analytique du bilinguisme chez l'enfant et de son apprentissage d'une langue seconde / Child Bilingualism and Second Language Learning: a descriptive bibliography.* Québec, 1975, 401 p.

F-5 GUNAR, Daniel. *Contact des langues et bilinguisme en Europe orientale: bibliographie analytique / Language Contact and Bilingualism in Eastern Europe: analytical bibliography.* Québec, 1979, 391 p.

*Epuisé / Out of print

Adresses des distributeurs / Distributors' addresses

Séries A, E, F:

PRESSES DE L'UNIVERSITE LAVAL,
C.P. 2447,
Québec, Québec,
Canada, G1K 7R4

INTERNATIONAL SCHOLARLY BOOK SERVICES INC.,
P.O. Box 555,
Forest Grove,
Oregon 97116, USA

CLUF/L'ECOLE,
11, rue de Sèvres,
75006 Paris,
France

Série B:

CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME,
Pavillon Casault, 6e sud,
Université Laval,
Québec, Québec,
Canada G1K 7P4

B-40, B-44, B-62, B-80:

ASSOCIATION CANADIENNE DE LINGUISTIQUE APPLIQUEE,
Institut des langues vivantes,
Université d'Ottawa,
59 est, avenue Laurier,
Ottawa, Ontario,
Canada, K1N 6N5

C-1, C-3, C-6:

MARCEL DIDIER LIMITEE,
2050, rue Bleury, suite 500,
Montréal, Québec,
Canada, H3A 2J4

C-2:

HARVEST HOUSE LIMITED,
4795 ouest, rue Sainte-Catherine,
Montréal, Québec,
Canada, H3Z 2B9

C-4:

INSTITUT DE L'UNESCO POUR L'EDUCATION,
Feldbrunnenstrasse 70,
Hambourg 13,
West Germany

APPROVISIONNEMENTS ET SERVICES CANADA,
Ottawa, Ontario,
Canada, K1A 0S9

C-5:

CENTER FOR APPLIED LINGUISTICS,
1611 North Kent Street,
Arlington,
Virginia 22209, USA

C-7, C-100, C-101, C-102, C-103, C-104, C-105, C-106, C-107:
NEWBURY HOUSE PUBLISHERS,
68 Middle Road,
Rowley,
Massachusetts 01969, USA

DIDACTA,
3465, Côte-des-Neiges, suite 61,
Montréal, Québec,
Canada, H3H 1T7

C-8:
MAISON DES SCIENCES DE L'HOMME,
ILTAM,
Esplanade des Antilles,
Domaine universitaire,
33405 Talence,
France

CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME,
Pavillon Casault, 6e sud,
Université Laval,
Québec, Québec,
Canada, G1K 7P4

C-9:
LIBRAIRIE KLINCKSIECK,
11, rue de Lille,
75007 Paris,
France

C-10:
TEXAS WESTERN PRESS,
University of Texas,
El Paso,
Texas 79968, USA

C-11:
MOUTON PUBLISHERS,
Noordeinde 41
2514 GC La Haye
(Netherlands)